Task Dynamic Coordination of the Speech Articulators: A Preliminary Model*

558

E. Saltzman
Haskins Laboratories, 270 Crown Street, New Haven, CT 06510, USA

INTRODUCTION

It is perhaps a truism that skilled actions of the limbs and speech articulators are goal directed. It is equally true, however, that such actions are performed by effector systems that are indifferent to the goals of would-be performers. An effector system is the set of limb segments or speech articulators used in a given action; a terminal device or end-effector is the part or a controlled effector system that is directly related to the goal of a performed action. Thus, in a reaching task, the fingers define the terminal device and the arm and hand comprise the effector system; in a "cup-to-mouth" task, the grasped cup is the terminal device and the combination of hand and arm constitutes the effector system; in a steady-state vowel production task, the tongue body is the terminal device and the jaw and tongue comprise the effector system. During skilled actions, the numerous degrees of freedom defined by the muscles and joints of such effector systems must be harnessed functionally in a manner specific to the task or goal at hand.

In addition to a skill's goal directedness, it is also clear that ordinary actions (such as walking or talking) or extraordinary actions (such as ballet or operatic singing) are never performed twice in exactly the same way. Yet observers and students of such activities seem to share the intuition that there is a task-specific commonality or invariance that underlies the separate task performances. In the present paper, a theoretical approach to these dual issues of contextual variation and task-specific invariance in skilled actions is described. This approach is called *task dynamics* (Saltzman and Kelso 1983a), and promises to provide within a single framework a parsimonious account of both variable and invariant aspects of well-learned, skilled actions. Saltzman and Kelso (1983a) describe how a mathematical, task-dynamic model can be applied to tasks involving relatively simple arm movements in the horizontal or sagittal planes (e.g., reaching discretely and cyclically, transporting cup-to-mouth, and crank-turning). The present paper describes how task-dynamic modeling is being extended by this author and his colleagues at Haskins Laboratories to the coordination and regulation of the speech articulators during linguistically meaningful tasks (cf. Browman and Goldstein 1985; Browman et al. 1984; Kelso et al. 1985).

There are (at least) two signature properties of skilled actions – *trajectory shaping* and *immediate compensation* – that must be accounted for by a theory of coordination and control. Trajectory

---

* This research was supported in part by NIH grant NS-13617 and contract no. N0014-83-K-0083 from the US Office of Naval Research.

Experimental Brain Research Series 15

shaping refers to the tendency of end-effector trajectories to display forms that are characteristic of the demands of performed tasks. For example, it has been demonstrated in several laboratories that in planar reaching tasks using the shoulder and elbow joints, the hand moves in a quasistraight line toward the target (e.g., Bizzi and Abend 1982; Bizzi et al. 1981; Morasso 1981; Soechting and Lacquaniti 1981; Wadman et al. 1980; see also Hollerbach and Atkeson, this volume). Similarly, in cup-to-mouth tasks, the grasped cup must maintain a spillage-preventing horizontal orientation while en route from table to mouth.

The second characteristic of skilled gestures, immediate compensation, refers to the task-specific flexibility of action systems in reorganizing themselves when faced with unexpected disturbances or perturbations. Thus, compensation for the perturbation of a given effector during a movement trajectory is achieved by readjusting the activity over the entire system in order to achieve the task goal (e.g., Bernstein 1967; Marsden et al. 1983; Nashner and McCollum 1985). Further, these readjustments appear to occur automatically without the need to detect the disturbance explicitly, replan a new movement, and execute the new movement plan. Kelso et al. (1984) have demonstrated such behavior in the speech articulators (jaw, upper and lower lip, tongue body) when subjects produced the utterances /baeb/ or /baez/ across a series of trials in which the jaw was occasionally and unpredictably tugged downward while moving upward to the final /b/ or /z/ constriction (see also, Abbs and Gracco 1983; Folkins and Abbs 1975). The system's response to the jaw perturbation was measured by observing the motions of the jaw and upper and lower lips as well as the electromyographic (EMG) activities of the orbicularis oris superior (upper lip), orbicularis oris inferior (lower lip), and genioglossus (tongue body) muscles. The investigators found relatively "immediate" task-specific compensation (i.e., 20-30 ms from onset of jaw pull to onset of compensatory response) in remote articulators to jaw perturbation. For /baeb/ (in which final lip closure is crucial) they found increased upper lip activity (motion and EMG) relative to the unperturbed control trials, but normal tongue activity; for /baez/ (in which final tongue-palate constriction is important) they found increased tongue activity relative to controls, but normal upper lip motion. The speed of these task-specific patterns suggests that compensation does not occur according to traditionally defined "intentional" reaction time processes, but rather according to an automatic, "reflexive" type of organization. However, such an organization is not defined in a hard-wired input/output manner. Instead, these data imply the existence of a selective pattern of coupling or gating among the component articulators that is specific to the utterance produced. Such compensatory behavior represents the classic phenomenon of motor equivalence (Hebb 1949; Lashley 1930), according to which a system will find alternate routes to a given goal if an initially intended route is unexpectedly blocked.

What type of coordinative processes could generate, in a task-specific manner, both characteristic trajectory patterns for unperturbed movements and spontaneous, compensatory behaviors for perturbed movements? The task-dynamic model for effector systems having many articulatory degrees of freedom was developed in an effort to deal with these issues (Saltzman and Kelso 1983a; see also, Boylls and Greene 1984, for related discussions of task-specific dynamics). The model is labeled task dynamic since: (a) it deals with the performance of well-learned skilled movements or gestures designed to accomplish real-world tasks; and (b) it is defined with respect to the dynamics that underlie a given action's kinematics. Note that kinematics refers to a gesture's observable spatiotemporal properties (e.g., its position, velocity, and acceleration trajectories over time), while dynamics refers to the pattern of the underlying field of forces that gives rise to these kinematics. The task-dynamic approach extends and elaborates the view that the functional units of

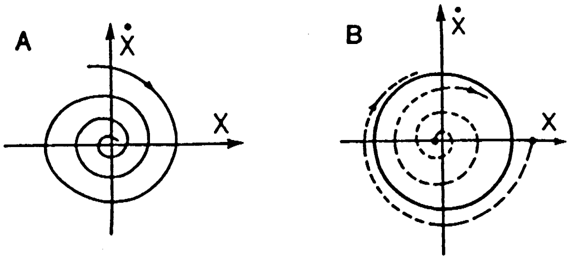

Fig. 1. Representative phase plane (x, $\dot{x}$) trajectories for point attractor (A) and periodic attractor (B) systems. Examples of motion equations are: A. $m\ddot{x} + b\dot{x} + kx = 0$, where m is mass, b is damping, and k is stiffness. B. $m\ddot{x} + b\dot{x} + kx = f(x,\dot{x})$, where $f(x,\dot{x})$ is a nonlinear damping (i.e., escapement) term, and all other coefficients are as in A

action (or *coordinative structures*; e.g., Easton 1972; Fowler 1977; Kelso et al. 1979; Turvey 1977) underlying the performance of a given gesture may be identified with abstractly defined, task–specific control regimes whose dynamic parameters (e.g., stiffness, damping, rest position) remain constant over the course of the gesture (cf. Fitch and Turvey 1978; Kelso et al. 1980; Kugler et al. 1980, 1982; Saltzman and Kelso 1985). In the task–dynamic model, the control regime that governs the performance of a particular gesture or task is defined functionally as an abstract (*task space*) dynamic system that is effector–independent, i.e., it does not explicitly incorporate the particular end–effectors directly involved in performing the task. It is hypothesized that a common task–space description underlies the functional equivalence of different effector systems for the performance of a given task, e.g., writing one's signature using a pencil held in the hand or between the teeth. Relatedly, qualitative differences between tasks are captured by corresponding *topological* distinctions among task–space dynamical systems (see also, Arbib 1984, for a related discussion of the relation between task and controller structures.)

For example, gestures involving a hand's discrete motion to a single spatial target and repetitive cyclic motion between two such targets are characterized by *point attractor* and *periodic attractor* dynamic regimes, respectively (cf. Abraham and Shaw 1982). The behaviors of these two types of dynamical systems may be represented in the phase plane (i.e., where system velocity is plotted vs position) as illustrated in Fig. 1, along with examples of corresponding equations of motion. Figure 1A shows a point attractor regime characterized by an (underdamped) mass–spring equation of motion. This system displays *point stability* or *equifinality*, in that it will asymptotically attain the equilibrium position, x_0, regardless of initial conditions for x and $\dot{x}$ and despite any transient perturbations encountered during its motion trajectory. Figure 1B shows a periodic attractor regime with a stable cyclic orbit (i.e., *limit cycle*) that is approached asymptotically by all trajectories (except those starting exactly at x_0) regardless of transiently introduced perturbations. The value of specifying a system's behavior in terms of topologically defined attractors is that such attractors provide task–specific, low dimensional descriptions for movement systems with many degrees of freedom, and promise to provide an elegant notational scheme for capturing the dynamical invariance across different effector systems that are observed to perform identical tasks. Distinct topologies correspond, therefore, to distinct patterns of task–dynamic parameters (e.g., damping and stiffness coefficients), and have been labeled the *organizational invariants* for skilled actions of different types (Fowler and Turvey 1978; Saltzman and Kelso 1983a, b). Such patterns denote functions that are preserved invariantly over changes in the parameter's specific values. In the task–dynamic model, the values of these *tuning* parameters (e.g., Greene 1972; Saltzman and Kelso 1983a, b) are determined according to factors such as the rate or amplitude of movement, and are defined to be constant over the course of a given gesture.

The task–dynamic model is able to account for the phenomena of trajectory shaping and immediate compensation without the need for explicit trajectory planning or replanning (see Saltzman and Kelso 1983a, for further details). Note that defining invariant patterns of dynamic parameters at the level of articulatory degrees of freedom (e.g., stiffness and damping parameters at the joints of an arm) will not suffice to generate these behaviors. Constant articulatory–dynamic parameters will not suffice to generate the quasistraight–line hand trajectories seen in planar reaching tasks (Delatizky 1982; Hollerbach 1982); rather, such trajectory shapes must result from task–specific patterns of change in these parameters during the reaching gestures. Similarly, the immediate compensation data for speech described above (Kelso et al. 1984) could not be generated by a system with a constant *rest configuration* parameter (i.e., a vector whose components are constant rest positions for the lips and jaw). As shown in these data, when sustained perturbations were introduced during articulatory closing gestures, the system "automatically" achieved the same constriction goals as for unperturbed gestures, but with different final or rest configurations. Thus, both trajectory shaping and immediate compensation behaviors appear to result from the way that dynamic parameters at the articulatory level are constrained to change during a gesture in a context–dependent manner. In the task–dynamic model, such patterns of constraint originate in corresponding invariant patterns of dynamic parameters at the task–space level of description.

Example 1. Planar Reaching, Three Joints

Using, for illustrative purposes, a discrete reaching task in the horizontal plane with angular motion at the shoulder, elbow, and wrist joints, the operation of a given task–dynamic regime may be understood in the following way. First, the functional aspects of a reaching gesture are specified in a two–dimensional task space as an invariant point attractor (e.g., a two–dimensional damped mass–spring system; see Fig. 2A). These dynamics give rise to an evolving pattern of state–dependent "forces" exerted on an effector–independent terminal device (i.e., a *task mass*). In the task space, the reach target defines the origin of a Cartesian coordinate system, with axis t_1 ("reach" axis) defined along a line from the initial position of the task mass to the target, and axis t_2 ("normal" axis) defined normal to t_1. The equations of motion for this task–dynamic regime are described in matrix notation as follows:

$$M_T\ddot{t} + B_T\dot{t} + K_T t = 0, \text{ where} \qquad (1)$$

$$M_T = \begin{bmatrix} m_T & 0 \\ 0 & m_T \end{bmatrix}; \; B_T = \begin{bmatrix} b_{T1} & 0 \\ 0 & b_{T2} \end{bmatrix};$$

$$K_T = \begin{bmatrix} k_{T1} & 0 \\ 0 & k_{T2} \end{bmatrix};$$

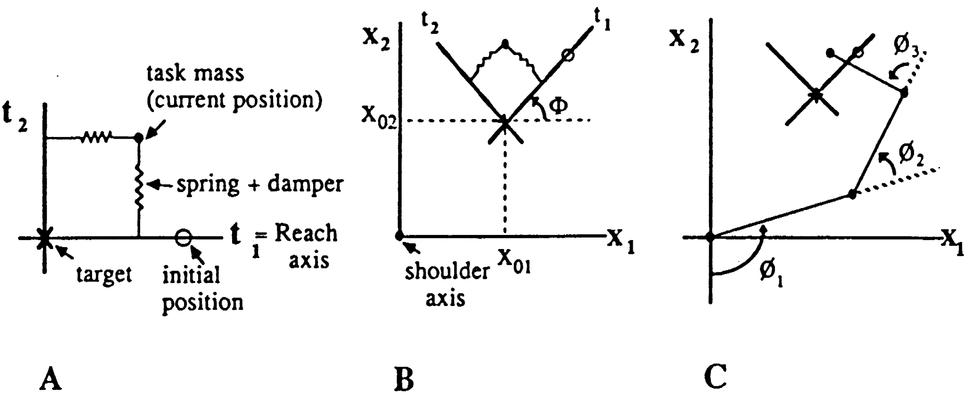

Fig. 2. Discrete reaching: A. Task space (t). B. Shoulder space (x). Task space is located and oriented in shoulder–centered reference frame via x_0 and Φ, respectively. C. Model articulator space (ϕ); ϕ's denote joint angles

where m_T is the task–mass coefficient; b_{T1}, b_{T2} are damping coefficients; and k_{T1}, k_{T2} are stiffness coefficients.

Equation (1) describes a linear, uncoupled set of task–space equations, whose terms are defined in units of force, and whose dynamic parameters (i.e., M_T, B_T, K_T) are constant. In Fig. 2A, the corresponding damping and stiffness elements are represented in lumped form by the squiggles in the lines connecting the task mass to axes t_1 and t_2.

Second, the task mass is identified with the relevant "virtual" end–effector (e.g., a virtual fingertip), and the task–space dynamic system is transformed kinematically into a two–dimensional body–space system (x_1, x_2; shoulder space) governing movements of the virtual end–effector (see Fig. 2B). Thus, the task space is located and oriented in body–space coordinates according to the tuning parameters $\mathbf{x}_0$ (the body–space position vector of the task–space origin) and Φ (the orientation angle between task axis t_1 and body axis x_1), respectively. The resulting set of linear body–space equations of motion for the task's terminal device are defined in matrix form as follows (in these and the following equations, a superscript T denotes the vector or matrix *transpose* operation):

$$M_B\ddot{\mathbf{x}} + B_B\dot{\mathbf{x}} + K_B\,\Delta\mathbf{x} = 0 \tag{2}$$

where $M_B = M_T R$, with M_T being the task–space mass matrix, and R the rotation transformation matrix with elements $r_{ij}(\Phi)$ converting task–space variables into body–space form; $B_B = B_T R$, where B_T is the task–space damping matrix; $K_B = K_T R$, where K_T is the task–space stiffness matrix; and $\Delta\mathbf{x}=\mathbf{x}-\mathbf{x}_0$, where $\mathbf{x}=(x_1, x_2)^T$, the current body–space position vector of the terminal device.

Note that Eq. (2), unlike Eq. (1), represents a set of body–space equations that are (usually) coupled due to the rotation transformation (i.e., the off–diagonal matrix elements are generally non-zero). However, as with the task-space equations, the terms of Eq. (2) are defined in force units and the resultant set of body-space dynamic parameters is constant.

Third, the body-space dynamic system is transformed into a three-dimensional "model" articulator space where the moving segments (upper arm, forearm, and hand) have lengths, but are massless (see Fig. 2C). Like the transformation from task space to body space, this transformation is a strictly kinematic one (since the segments have no mass) and involves only the substitution of variables defined in one coordinate system for variables defined in another coordinate system. As illustrated in Fig. 2C, this corresponds to expressing body-space variables ($\mathbf{x}$, $\dot{\mathbf{x}}$, $\ddot{\mathbf{x}}$) as functions of an arm model's kinematic variables ($\boldsymbol{\phi}$, $\dot{\boldsymbol{\phi}}$, $\ddot{\boldsymbol{\phi}}$; where $\boldsymbol{\phi} = [\phi_1, \phi_2, \phi_3]^T$, and ϕ_1 is the shoulder angle defined relative to axis x_2, ϕ_2 is the elbow angle defined relative to the upper arm segment, ϕ_3 is the wrist angle defined relative to the forearm segment), and the arm's proximal (shoulder) and distal (fingertip) ends are attached to the body-space origin and the terminal device/task mass, respectively. The body-space variables of Eq. (2) are transformed into the joint-angle variables of the massless arm model using the following kinematic relationships:

$$
\begin{aligned}
\mathbf{x} &= \mathbf{x}\,(\boldsymbol{\phi}) \\
\dot{\mathbf{x}} &= \mathrm{J}\,(\boldsymbol{\phi})\,\dot{\boldsymbol{\phi}} && \text{(3a)} \\
\ddot{\mathbf{x}} &= \mathrm{J}\,(\boldsymbol{\phi})\,\ddot{\boldsymbol{\phi}} + (\mathrm{dJ}\,(\boldsymbol{\phi})/\mathrm{dt})\dot{\boldsymbol{\phi}} && \text{(3b)} \\
&= \mathrm{J}\,(\boldsymbol{\phi})\ddot{\boldsymbol{\phi}} + \mathrm{V}(\boldsymbol{\phi})\,\dot{\boldsymbol{\phi}}_p && \text{(3c)}
\end{aligned}
$$

where $\mathbf{x}(\boldsymbol{\phi}) = (x_1(\boldsymbol{\phi}), x_2(\boldsymbol{\phi}))^T$, the current body-space position vector of the terminal device expressed as a function of the current model arm configuration; $\dot{\boldsymbol{\phi}}_p = [\dot{\phi}^2_1, \dot{\phi}_1\dot{\phi}_2, \dot{\phi}_1\dot{\phi}_3, \dot{\phi}^2_2, \dot{\phi}_2\dot{\phi}_3, \dot{\phi}^2_3]^T$ is the current model arm joint velocity product vector; $\mathrm{J}(\boldsymbol{\phi})$ is the *Jacobian* transformation matrix whose elements J_{ij} are partial derivatives, $\partial x_i/\partial\phi_j$, evaluated at the current $\boldsymbol{\phi}$; and $\mathrm{V}(\boldsymbol{\phi})$ is a matrix resulting from rearranging the terms of the expression $(\mathrm{dJ}(\boldsymbol{\phi})/\mathrm{dt})\,\dot{\boldsymbol{\phi}}$ in order to segregate the joint velocity products into a single vector $\dot{\boldsymbol{\phi}}_p$.

Using the kinematic relationships in Eq. (3), the model effector system's equation of motion is as follows:

$$
\mathrm{M_B}\mathrm{J}\ddot{\boldsymbol{\phi}} + \mathrm{B_B}\mathrm{J}\dot{\boldsymbol{\phi}} + \mathrm{K_B}\Delta\mathbf{x}(\boldsymbol{\phi}) = -\mathrm{M_B}\mathrm{V}\dot{\boldsymbol{\phi}}_p \tag{4}
$$

where $\mathrm{M_B}$, $\mathrm{B_B}$, $\mathrm{K_B}$ are the same constant matrices used in Eq. (2); and $\Delta\mathbf{x}(\boldsymbol{\phi}) = \mathbf{x}(\boldsymbol{\phi}) - \mathbf{x}_0$, where $\mathbf{x}_0$ is the same constant vector used in Eq. (2). It should be noted that since $\Delta\mathbf{x}$ in Eq. (2) and (4) is *not* assumed to be "small," a differential approximation $\mathbf{dx} = \mathrm{J}(\boldsymbol{\phi})\,\mathbf{d}\boldsymbol{\phi}$ is not justified and, therefore, Eq. (3a) was used instead for the kinematic displacement transformation into model arm variables.

The terms of Eq. (4) are still defined in units of force, not torque, and may be rewritten in units of angular acceleration:

$$
\ddot{\boldsymbol{\phi}} + \mathrm{J^*}\mathrm{M_B}^{-1}\mathrm{B_B}\mathrm{J}\dot{\boldsymbol{\phi}} + \mathrm{J^*}\mathrm{M_B}^{-1}\mathrm{K_B}\Delta\mathbf{x}(\boldsymbol{\phi}) + \mathrm{J^*}\mathrm{V}\dot{\boldsymbol{\phi}}_p = 0 \tag{5}
$$

where $\mathrm{J^*}$ is a weighted Jacobian pseudoinverse (e.g., Benati et al. 1980; Klein and Huang 1983; Whitney 1972) that is used because there are a greater number of model articulator variables than spatial variables for this task. Hence, the model effector system is *redundant* (e.g., Saltzman 1979), the *inverse kinematic* transform from spatial to model articulator coordinates is indeterminate, and the Jacobian inverse (J^{-1}) cannot be defined. More specifically, $J^* = A^{-1}J^T(JA^{-1}J^T)^{-1}$, where A is a positive definite *articulatory weighting* matrix whose elements are constant during a given gesture. Using J^* provides a unique, optimal least squares solution for the differential transformation from body-space to model articulator variables that is weighted according to the pattern of elements in the A-matrix. In current modeling, the A-matrix is defined to be of diagonal form, and a given set of articulator weights will constrain motion of an articulator in direct proportion to the magnitude of the corresponding weighting element. Hence, different articulator weighting patterns are associated with different patterns of relative angular motions of the three joints for the same task-space motion of the task mass (or body-space motion of the virtual fingertip). For example, one weighting pattern might correspond to predominant shoulder motion, while a second weighting pattern might correspond to predominant elbow motion for the same task- or body-space trajectory of the terminal device. In this sense, elements of the A-matrices used in the associated J^*'s define a further set of tuning parameters for the model effector system's equation of motion (Eq. 5).

The task-dynamic model allows one to define for the discrete reach (as well as other tasks) an invariant task-space dynamic regime that: (a) is specified by a constant set of task-dynamic parameters, and (b) constrains in a context-dependent way the evolving pattern of changes in the model arm's articulatory-dynamic parameters (i.e., stiffness, damping, and equilibrium positions of shoulder, elbow, and wrist joints) during the course of the gesture. Thus, one may interpret the task-specific, coherent movements of the model effector system as resulting from the way that instantaneous task-space "forces" acting on the associated terminal device are distributed across the model arm's articulatory degrees of freedom during the course of the planar reach. At any given instant during this gesture, the partitioning is based on two factors:

1. The task-specific, constant set of task space (Eq. 1), body space (Eq. 2), and model articulator space (Eq. 4 and 5) dynamic parameters
2. The current values of elements in the posturally dependent transformation matrices (i.e., the J and J^T matrices in Eq. 4 and 5) that relate motions of the articulators at their current configuration to corresponding body-space motions of the virtual fingertip. Because these elements are nonlinear functions of the current arm model posture ϕ, the elements of the matrix products in Eq. (4) and (5) (i.e., the coefficients that define articulatory-dynamic parameters) are also dependent on the evolving configuration of the arm model.

The final step in the task-dynamic approach is to exploit algebraic relations between the model arm's dynamic regime and the physical and control parameters of the "real" (biological, robotic, or prosthetic) arm in order to specify patterns of control parameters over time for the real arm. Saltzman and Kelso (1983a) discuss two related methods for specifying these controls. Both methods are applicable to the control and coordination of artificial linkage systems (e.g., robotic or prosthetic devices), although one offers a more biologically plausible style of control than the other (see also Hogan and Cotter 1982). The aim of both methods, however, is to make the real arm behave identically or near-identically to the model arm. Further, the essence of the task-dynamic approach lies in its account of the coordinated movement patterns that arise in a task-specific and posturally conditioned form in the model effector system. Consequently, for the purposes of the present paper, further discussion will focus on behavioral phenomena in the model articulators only. The interested reader is referred to Saltzman and Kelso (1983a), however, for details concerning the hypothesized relationships between control processes of the model and real effector systems.

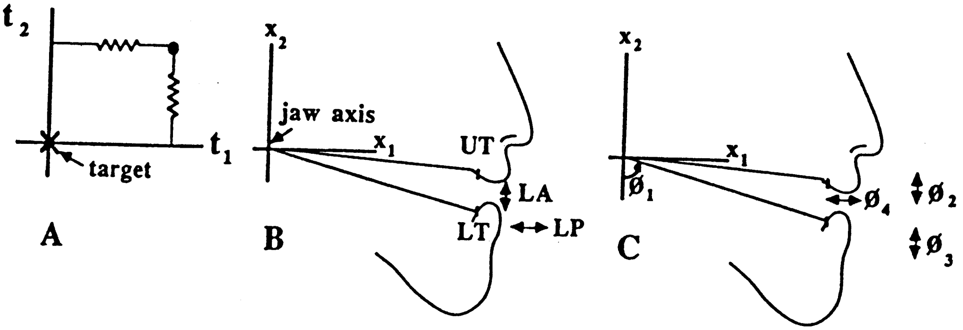

Fig. 3. Bilabial tasks: A. Task space (t). Closed circle denotes current system configuration. Squiggles denote each axis' dynamics in lumped forms. B. Jaw space (x). Local tract variables (LP, lip protrusion; LA, lip aperture) are expressed in jaw coordinates. UT and LT denote positions of upper and lower front teeth, respectively. C. Model articulator space (ϕ). ϕ's denote articulator variables

TASK DYNAMICS AND SPEECH: BILABIAL GESTURES

The task-dynamic approach has been extended in a preliminary way to speech gestures in order to explore the hypothesis that speech production involves task-specific, dynamically specified coordination of the articulators.

Example 2. Discrete Bilabial Closure, Unperturbed Gestures

As with the limb tasks described earlier, the first step in generating simulated movements of the speech articulators is to specify the functional aspects of these gestures with reference to the movements of an effector-independent terminal device (i.e., an idealized vocal tract constriction). This is done in a two-dimensional task space whose axes represent constriction location (t_1) and constriction degree (t_2), and the topological structure of the control regime for each task-space variable is specified according to the qualitative characteristics of the given speech task. Thus, for example, discrete and repetitive speech gestures will have point attractor and limit cycle regimes, respectively, along each axis. At the task-space level, then, the control regime is an abstract one in that the constriction being controlled is independent of any particular effector system, and can refer, for example, to either a bilabial constriction produced by the lips and jaw or to a tongue-palate constriction produced by the tongue and jaw. Since simulations to date have focused on bilabial gestures, we will begin by examining a discrete bilabial closure task involving (uncoupled) point attractor dynamics along each task axis (see Fig. 3A). The task-space equation of motion is expressed as follows:

$$M_T\ddot{t} + B_T\dot{t} + K_T t = 0 \tag{6}$$

where $M_T = \begin{bmatrix} m_{T1} & 0 \\ 0 & m_{T2} \end{bmatrix}$; $B_T = \begin{bmatrix} b_{T1} & 0 \\ 0 & b_{T2} \end{bmatrix}$; and $K_T = \begin{bmatrix} k_{T1} & 0 \\ 0 & k_{T2} \end{bmatrix}$

In Eq. (6) m_{T1}, m_{T2} are inertial coefficients; b_{T1}, b_{T2} are damping coefficients; and k_{T1}, k_{T2} are stiffness coefficients.

The forms of these task-space dynamics and corresponding equation of motion are identical to those for the discrete limb reaching task described earlier (Fig. 2A; Eq. 1). This identity highlights the fact that functional equivalence among tasks does not depend on the specific effector systems involved, but only on the topological equivalence of dynamical regimes in the task spaces. The two main differences between the limb and speech examples are: (a) the task-space axes for the bilabial task do not share a common task mass, but rather are characterized by their own inertial coefficients m_{T1} and m_{T2} (compare Eq. 2 and 6); and (b) the axes for the bilabial task are not differentiated into distinct "reach" and "normal" axes as they were in the limb reaching task. Finally, as in the reaching example, movements along the task axes do not influence one another, since the corresponding equations of motion are defined to be uncoupled.

The next step in modeling the bilabial closure is to transform the task-space system kinematically into a two-dimensional body-space system (x_1, x_2) defined in the midsagittal plane of the vocal tract and centered on the jaw's rotation axis (see Fig. 3B). In contrast to the task-space regime, the body-space dynamics are effector-specific, in that they refer to the movement of a "virtual" terminal device (i.e., the bilabial constriction) of the effector system defined by the lips and jaw. The result of transforming from task-space (t_1, t_2) to jaw-space (x_1, x_2) coordinates, then, is to define a two-dimensional set of motion equations, one for each axis of jaw space. As with the task-space equation, the jaw-space equation has the same form as its corresponding shoulder-space reaching equation (Eq. 2). The jaw-space equation is as follows:

$$M_B\ddot{\mathbf{x}} + B_B\dot{\mathbf{x}} + K_B\Delta\mathbf{x} = \mathbf{0} \tag{7}$$

where $\mathbf{x}, \dot{\mathbf{x}}, \ddot{\mathbf{x}} = (x_1, x_2)^T$ and its derivatives with respect to time; $\Delta\mathbf{x} = \mathbf{x} - \mathbf{x}_0$, where $\mathbf{x}_0$ is the target vector for lip protrusion (x_{01}) and lip aperture (x_{02}); $M_B = M_T$, $B_B = B_T$, and $K_B = K_T$, since no rotation is involved in the transformation from task- to jaw-space coordinates.

Equation (7) contains a constant set of dynamic parameters, and governs the movements for the bilabial constriction along the dimensions of lip aperture (LA) and lip protrusion (LP). Lip aperture and protrusion are labeled *local tract variables*, and represent the effector-specific body-space versions of the effector-independent task-space variables of constriction degree and location, respectively. Lip aperture is defined by the vertical distance between the upper and lower lips, and lip protrusion by the horizontal distances in the anterior-posterior direction of the upper and lower lips from the upper and lower teeth, respectively. It should be noted that upper and lower lip protrusion movements are not independent in this formulation, but have been constrained to be equal in the model for purposes of simplicity. Consequently, like constriction location in task space, lip protrusion in body space constitutes only a single degree of freedom. Finally, it should be noted that the control regimes for each jaw-space coordinate are independent, since their corresponding equations of motion are uncoupled. This is due to the fact that lip aperture and protrusion are defined parallel to the x_2 and x_1 jaw-space axes, respectively. Note, however, that such non-interacting dynamics are not usually found at the body-space level of description. For example, with movements of the tongue body orthogonal and tangential to the (curved) palate, the set of uncoupled task-space equations would be transformed into a set of (generally) coupled jaw-space equations.

The last step in modeling the closure is to transform kinematically the two-dimensional jaw-space regime into the coordinates of a four-dimensional model articulator space. The model articulators are moving segments that have lengths but are massless (see Fig. 3C), and are defined with reference to the simplified articulatory degrees of freedom adopted in the Haskins Laboratories software articulatory speech synthesizer (Rubin et al. 1981). For bilabial gestures, the articulator set associated with lip aperture includes rotation of the jaw ($\mathring{a}_1$), and vertical displacements of the upper lip ($\mathring{a}_2$) and lower lip ($\mathring{a}_3$) relative to the upper and lower front teeth, respectively; for lip protrusion, the articulator set includes yoked horizontal displacements in the anterior-posterior direction of the upper and lower lips ($\mathring{a}_4$) relative to the upper and lower front teeth, respectively. Expressed in units of linear acceleration, the model articulator equation has the same form as Eq. 5 and is expressed as follows (the angular acceleration terms in the jaw's motion equation have been multiplied by a unit scaling factor to ensure dimensional homogeneity along all articulatory degrees of freedom):

$$\ddot{\phi} + J^*{M_B}^{-1}B_BJ\dot{\phi} + J^*{M_B}^{-1}K_B\Delta x(\phi) + J^*V\dot{\phi}_p = 0 \tag{8}$$

where M_B, B_B, K_B are the same constant matrices used in Eq. (7); and $\Delta x(\phi) = x(\phi) - x_0$, where $x(\phi)$ is expressed as a function of model articulator variables, and x_0 is the same constant vector used in Eq. (7); the elements of the Jacobian matrix (J, and hence also V and J*) reflect the geometrical relationship among motions of the (simplified) model speech articulators (4 degrees of freedom) and motions of the corresponding local tract variables (2 degrees of freedom); and A is the articulator weighting matrix, a component of the pseudoinverse J*. A's elements reflect task-specific constraints on the relative motions of the articulators during the closing gesture.

Given a fixed set of tuning parameters (i.e., M_T, B_T, K_T, x_0, and A) and a set of initial conditions (ϕ_I, $\dot{\phi}_I$, and hence a corresponding x_I and $\dot{x}_I$) Eq. 8 will generate a pattern of coordinated motion in the model speech articulators that will achieve the task goals specified for the local tract variables. For an initial configuration (ϕ_I) corresponding to open and relatively unprotruded lips, and with an initial velocity vector of zero, the coordinated articulator movements will reflect the evolving task-specific motions of the local tract variables en route to their specified targets (x_0), with motion characteristics (e.g., speed, degree of overshoot, etc.) specified by the pattern of M_T, B_T, and K_T parameters. Assuming the system is not perturbed during its motion trajectory, the relative extents of movement for the articulators associated with lip aperture (i.e., ϕ_1, ϕ_2, ϕ_3 in Fig. 3C) will be specified by the relative values of articulator weights in the associated articulatory weighting matrix, A. Figure 4 (configurations A and B) illustrates an unperturbed movement from an initially open and relatively unprotruded configuration (Fig. 4A) to a closed and relatively protruded final configuration (Fig. 4B). Since the articulators associated with lip aperture were weighted equally in the corresponding A-matrix, the extents of motion for these articulators were equal over the course of the gesture.

Example 3. Immediate Compensation, Bilabial Closure, Perturbed Gestures

Previous dynamical accounts of coordinated actions performed by the limbs and speech articulators have posited that invariant sets of dynamic parameters could be defined at the level of articulatory

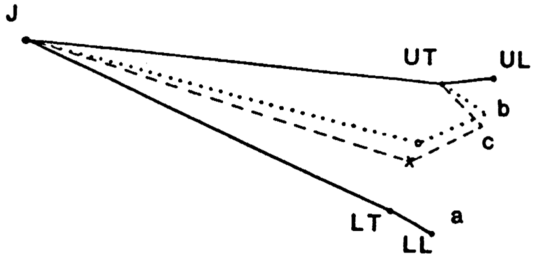

Fig. 4. Simulated articulator configurations for bilabial closure task. A. Initial configuration (solid lines). B. Final configuration, unperturbed trajectory (dotted lines). C. Final configuration, perturbed trajectory (broken lines). Note that closure occurs lower in jaw space in C than in B. J, jaw axis; UT, upper teeth; UL, upper lip; LT, lower teeth; LL, lower lip

degrees of freedom (e.g., Cooke 1980; Fel'dman 1966; Fowler 1977; Kelso 1977; Polit and Bizzi 1978). Thus, for example, discrete targeting tasks of the elbow joint were modeled as damped mass–spring systems (having point attractor dynamics) where the target angle was specified by the value of the rest angle dynamic parameter. As discussed earlier, this approach implies that the task of reaching a bilabial closure target for speech is specified according to a corresponding rest–configuration parameter for the articulators. However, recent work (Abbs and Gracco 1983; Folkins and Abbs 1975; Kelso et al. 1984) has shown that this formulation must be modified. In particular, the Kelso et al. (1984) study demonstrated that if the jaw is retarded en route to a bilabial closure target for /b/, then the closure is still attained and the final articulatory configuration for the perturbed movement is different from the final configuration for unperturbed movements. Significantly, the upper lip compensation is absent if the jaw is perturbed en route to an alveolar closure target for /z/. These results show that an invariant dynamic description of a movement does not apply at the articulator level, since the articulatory–dynamic parameters must be able to change according to a movement's context in an utterance–specific (i.e., /b/ vs /z/) manner. Furthermore, the speed of these compensatory behaviors suggests that they must occur "automatically" without reference to traditional stimulus–response reaction–time correction procedures.

The task–dynamic model handles such immediate compensation as follows. Bilabial closing gestures are simulated as discrete movements toward target constrictions, using point attractor dynamics for the local tract variables of lip aperture and protrusion (see Eq. 7 above). When the simulated jaw is "frozen" in place during the closing gesture at the level of the model effector system, the main qualitative features of the perturbation data are captured, in that: (a) compensation is immediate in the upper and lower lips to the jaw perturbation, i.e., the system does not require reparameterization in order to compensate; and (b) the target bilabial closure is reached (although with different final articulator configurations and, hence, different jaw–space locations for the closure) for both perturbed (Fig. 4C) and unperturbed (Fig. 4B) "trials."

Example 4. Cyclic Bilabial Motion, Unperturbed Gestures

The point attractor task–space (and local tract variable) topology that was used in the discrete bilabial closure task is inappropriate, however, for generating cyclic bilabial gestures, e.g., a

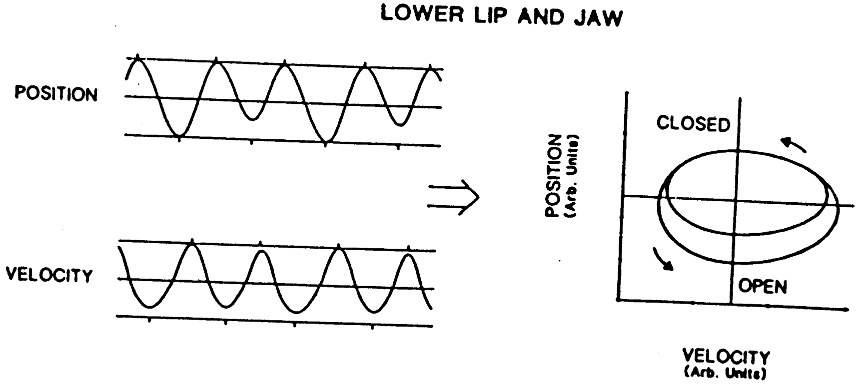
LOWER LIP AND JAW

Fig. 5. Simulated trajectories for lower lip height (i.e., jaw and lower lip) in the time domain (left) and phase plane (right) for a repetitive sequence of /ma/'s with alternating stress (from Kelso et al. 1985)

sequence of repeated /ma/'s as in "...mamama...." The task-dynamic model has been used to simulate a repetitive gestural sequence that is characterized by an alternating stress pattern, e.g., "...mamamama...," where the underlining denotes the pattern of stress (Browman et al. 1984). Mass-spring dynamics were specified for the local tract variables of lip aperture and protrusion in order to generate sustained cyclic motions of the model articulators. (Mass-spring and limit cycle dynamics can produce near-identical motion trajectories in the absence of perturbations to the system. Since the modeled cyclic bilabial gestures were unperturbed by design, (undamped) mass-spring dynamics were adequate for these purposes. The model is presently being extended, however, to include limit cycle dynamics at the task-space level, in order to explore the simulated effects of perturbations introduced during cyclic speech tasks.) Focusing on lip aperture, the parameters of rest position and stiffness were estimated from articulatory movement data collected in an experiment on reiterantly produced speech (Kelso et al. 1985). In reiterant speech, talkers substitute a given syllable (e.g., /ma/) for the real syllables in an utterance while maintaining the utterance's normal stress pattern (e.g., the sentence "When the sunlight strikes raindrops in the air" becomes "ma ma ma ma ma ma ma ma ma ma"). The lip aperture parameters for the task-dynamic simulation were estimated using the average amplitudes and frequencies of the articulatory data obtained for the stressed and unstressed syllables spoken reiterantly at a given rate. Figure 5 illustrates the resultant cyclic trajectories for lower lip height, both in the time domain and the phase plane. For a given simulated cyclic gesture (closure-to-closure), the equilibrium position was set only once because, in the data, the jaw-lip complex returned roughly to the same position at closure for each syllable. The values for the equilibrium positions in temporally adjacent cycles alternated in value, however, since stressed syllables were found to involve greater movement amplitudes than unstressed syllables. Additionally, because closing gestures were faster than opening gestures in these data, two values of stiffness were specified within each cycle: one at the start of the opening gesture and another at the start of the closing gesture. The set of task-dynamic parameters were invariant, therefore, over the course of a given opening or closing gesture.

CONCLUSION

The task–dynamic model is able to generate coordinated movement patterns for the model articulators in both discrete and cyclic unperturbed (bilabial) utterances. Additionally, for discrete bilabial closing gestures it provides task–specific patterns of compensatory responses to jaw perturbations that are qualitatively similar to those observed experimentally. Finally, Browman et al. (1984) have used sets of simulated articulator trajectories from an alternating stress, repetitive, bilabial speech task as inputs to the Haskins Laboratories articulatory speech synthesizer (Rubin et al. 1981; see also Example 2 above) with promising acoustic and perceptual results. Note that, although these simulated utterances involve a simple stress pattern and segmental structure, the task–dynamic approach to articulatory speech synthesis could certainly be used to generate more complex utterances on a gesture–by–gesture basis. The elegance of the procedure would still be maintained, however, since utterance–specific and contextually variable patterns of articulator trajectories and compensatory responses would still emerge automatically as implicit consequences of task space control regimes that are invariant within a given speech gesture. There is no need to invoke either explicit trajectory planning or replanning procedures on a timeframe–to–timeframe basis within the gesture. My colleagues and I are encouraged by these preliminary results, and are currently engaged in extending the task–dynamic model to account for phenomena such as coarticulation (e.g., Harris 1984; Kent and Minifie 1977) and relative timing (e.g., Kent et al. 1974; Tuller et al. 1982, 1983) among serially ordered speech gestures.

REFERENCES

Abbs JH, Gracco VL (1983) Sensorimotor actions in the control of multimovement speech gestures. Trends Neurosci 6: 393–395

Abraham RH, Shaw CD (1982) Dynamics – the geometry of behavior. Aerial, Santa Cruz

Arbib MA (1984) From synergies and embryos to motor schemas. In: Whiting HTA (ed) Human motor actions: Bernstein reassessed. North–Holland, Amsterdam

Benati M, Gaglio S, Morasso P, Tagliasco V, Zaccaria R (1980) Anthropomorphic robotics. II. Analysis of manipulator dynamics and the output motor impedance. Biol Cybern 38: 141–150

Bernstein NA (1967) The coordination and regulation of movements. Pergamon, London

Bizzi E, Abend W (1982) Posture control and trajectory formation in single and multiple joint arm movements. In: Desmedt JE (ed) Brain and spinal mechanisms of movement control in man: new developments and clinical applications. Raven, New York

Bizzi E, Accornero N, Chapple W, Hogan N (1981) Processes underlying arm trajectory formation. In: Ajmone–Marsan C, Pompeiano O (eds) Brain mechanisms of perceptual awareness and purposeful behavior. IBRO monograph series. Raven, New York, pp 311–318

Boylls CC Jr, Greene PH (1984) Introduction: Bernstein's significance today. In: Whiting HTA (ed) Human motor actions: Bernstein reassessed. North–Holland, Amsterdam

Browman CP, Goldstein LM (1985) Dynamic modeling of phonetic structure. In: Fromkin V (ed) Phonetic linguistics. Academic, New York

Browman CP, Goldstein L, Kelso JAS, Rubin P, Saltzman E (1984) Articulatory synthesis from underlying dynamics. J Acoust Soc Am 75: S22 – S23 (abstract)

Cooke JD (1980) The organization of simple, skilled movements. In: Stelmach GE, Requin J (eds) Tutorials in motor behavior. North – Holland, Amsterdam

Delatizky J (1982) Final position control in simulated planar horizontal arm movements. PhD thesis, Department of Electrical Engineering and Computer Science, MIT

Easton TA (1972) On the normal use of reflexes. Am Sci 60: 591 – 599

Fel'dman AG (1966) Functional tuning of the nervous system with control of movement or maintenance of a steady posture. III. Mechanographic analysis of execution by man of the simplest motor tasks. Biophysics 11: 766 – 775

Fitch H, Turvey MT (1978) On the control of activity: some remarks from an ecological point of view. In: Landers D, Christina R (eds) Psychology of motor behavior and sports. Human Kinetics, Urbana

Folkins JW, Abbs JH (1975) Lip and jaw motor control during speech: responses to resistive loading of the jaw. J Speech Hear Res 18: 207 – 220

Fowler C (1977) Timing control in speech production. Indiana University Linguistics Club, Bloomington

Fowler CA, Turvey MT (1978) Skill acquisition: an event approach with special reference to searching for the optimum of a function of several variables. In: Stelmach G (ed) Information processing in motor control and learning. Academic, New York

Greene PH (1972) Problems of organization of motor systems. Prog Theoret Biol 2: 303 – 338

Harris KS (1984) Coarticulation as a component of articulatory descriptions. In: Daniloff RG (ed) Articulation assessment and treatment issues. College Hill, San Diego

Hebb DO (1949) The organization of behavior. Wiley, New York

Hogan N, Cotter SL (1982) Cartesian impedance control of a nonlinear manipulator. In: Book WJ (ed) Robotics research and advanced applications. ASME, New York, pp 121 – 128

Hollerbach JM (1982) Computers, brains, and the control of movement. Trends Neurosci 5: 189 – 192

Kelso JAS (1977) Motor control mechanisms underlying human movement reproduction. J Exp Psychol: Hum Percept Perform 3: 529 – 543

Kelso JAS, Southard DL, Goodman D (1979) On the nature of human interlimb coordination. Science 203: 1029 – 1031

Kelso JAS, Holt KG, Kugler PN, Turvey MT (1980) On the concept of coordinative structures as dissipative structures: II. Empirical lines of convergence. In: Stelmach G, Requin J (eds) Tutorials in motor behavior. North – Holland, Amsterdam, pp 49 – 70

Kelso JAS, Tuller B, V – Bateson E, Fowler CA (1984) Functionally specific articulatory cooperation following jaw perturbations during speech: evidence for coordinative structures. J Exp Psychol: Hum Percept Perform 10: 812 – 832

Kelso JAS, V – Bateson E, Saltzman EL, Kay BA (1985) A qualitative dynamic analysis of reiterant speech production: phase portraits, kinematics, and dynamic modeling. J Acoust Soc Am 77: 266 – 280

Kent RD, Minifie FD (1977) Coarticulation in recent speech production models. J Phon 5: 115 – 133

Kent RD, Carney PJ, Severeid LR (1974) Velar movement and timing: evaluation of a model for binary control. J Speech Hear Res 17: 470 – 488

Klein CA, Huang C – H (1983) Review of pseudoinverse control for use with kinematically redundant manipulators. IEEE Trans Syst Man Cyber SMC – 13: 245 – 250

Kugler PN, Kelso JAS, Turvey MT (1980) On the concept of coordinative structures as dissipative structures: I. Theoretical lines of convergence. In: Stelmach GE, Requin J (eds) Tutorials in motor behavior. North – Holland, Amsterdam, pp 3 – 47

Kugler PN, Kelso JAS, Turvey MT (1982) On the control and coordination of naturally developing systems. In: Kelso JAS, Clark JE (eds) The development of movement control and coordination. Wiley, Chichester, pp 5 – 78

Lashley KS (1930) Basic neural mechanisms in behavior. Psychol Rev 37: 1 – 24

Marsden CD, Merton PA, Morton HB (1983) Rapid postural reactions to mechanical displacement of the hand in man. In: Desmedt JE (ed) Motor control mechanisms in health and disease. Raven, New York

Morasso P (1981) Spatial control of arm movements. Exp Brain Res 42: 223 – 227

Nashner LM, McCollum G (1985) The organization of human postural movements: a formal basis and experimental synthesis. Behav Brain Sci 8: 135 – 172

Polit A, Bizzi E (1978) Processes controlling arm movements in monkeys. Science 201: 1235 – 1237

Rubin P, Baer T, Mermelstein P (1981) An articulatory synthesizer for perceptual research. J Acoust Soc Am 70: 321–328

Saltzman E (1979) Levels of sensorimotor representation. J Math Psychol 20: 91–163

Saltzman E, Kelso JAS (1983a) Skilled actions: a task dynamic approach. Hask Lab Status Rep Speech Res SR–76: 3–50

Saltzman E, Kelso JAS (1983b) Toward a dynamical account of motor memory and control. In: Magill R (ed) Memory and control of action. North–Holland, Amsterdam, pp 17–38

Saltzman E, Kelso JAS (1985) Synergies: stabilities, instabilities, and modes. Behav Brain Sci 8: 161–163

Soechting JE, Lacquaniti F (1981) Invariant characteristics of a pointing movement in man. J Neurosci 1: 710–720

Tuller B, Kelso JAS, Harris KS (1982) Interarticular phasing as an index of temporal regularity in speech. J Exp Psychol: Hum Percept Perform 8: 460–472

Tuller B, Kelso JAS, Harris KS (1983) Further evidence for the role of relative timing in speech: a reply to Berry. J Exp Psychol: Hum Percept Perform 9: 829–833

Turvey MT (1977) Preliminaries to a theory of action with reference to vision. In: Shaw R, Bransford J (eds) Perceiving, acting and knowing: toward an ecological psychology. Erlbaum, Hillsdale

Wadman WJ, Denier van der Gon JJ, Derksen RJA (1980) Muscle activation patterns for fast goal–directed arm movements. J Human Movement Stud 6: 19–37

Whitney DE (1972) The mathematics of coordinated control of prosthetic arms and manipulators. ASME J Dyn Sys Measurement Control 94: 303–309